UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/727,169 | 12/02/2003 | Michael L. Kazar | SPIN-5 | 5977 |

7590 05/04/2006

Ansel M. Schwartz
Suite 304
201 N. Craig Street
Pittsburgh, PA 15213

| EXAMINER | |
|---|---|
| DOAN, DUC T | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2188 | |

DATE MAILED: 05/04/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/727,169 | KAZAR ET AL. | |
| | Examiner | Art Unit | |
| | Duc T. Doan | 2188 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *03 April 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-39* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *36 and 37* is/are allowed.

6)☒ Claim(s) *1-35,38-3* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

## DETAILED ACTION

### *Status of Claims*

Claims 1-39 have been presented for examination in this application. In response to the last Office Action, none of claims have been amended. As a result, claims 1-39 are now pending in this application.

Applicant's arguments filed 4/3/06 have been fully considered with the results as follows:

Claims 1-35,38-39 are rejected.

Claims 36-37 are allowed.

### *Claim Objections*

Claims 1,38 is objected to because of the following informalities:

As per claim 1, the NFS should not be abbreviated for the initial recital in the claims.

As per claim 31, the VFS should not be abbreviated for the initial recital in the claims.

All dependent claims are objected to as having the same deficiencies as the claims they depend from.

Appropriate correction is required.

### *Claim Rejections - 35 USC § 103*

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-11 rejected under 35 U.S.C. 103(a) as being unpatentable over Fridella et al (US Pub 2005/0044090) in view of APA (US application 10/727169) and further in view of Edsall et al (US 2004/0139167).

As for claim 1, Frieda describes an apparatus for data storage comprising: a cluster of NFS servers, each server having network ports for incoming file system requests and cluster traffic between servers (Fig 2: #133, page 1 paragraph 2); and a plurality of storage arrays in communication with the servers (Fig 2: #135), the servers utilizing a striped file system for storing data. Frieda does not describe the claim's aspect of a striped file system. However APA describes a parallel virtual file system capable of accessing data concurrently vie multiple servers. It would have been obvious to one of ordinary skill in the art at the time of invention to include the parallel virtual file system as suggested by APA in Frieda's system to concurrently accessing secondary storage in a parallel manner, thereby the system bandwidth can be scaled accordingly by providing multiple servers (APA's page 1).

As for claim 2, the claim recites wherein each server has a network element and a disk element. Claim 2 rejected based on the same rationale as in the rejection of claim 1.

As for claims 3-4, the claim recites; wherein each disk element has a virtual file system with the virtual file system of each disk element together forming a striped VFS (claim 3);

wherein all disk elements for a virtual file system act as meta-data servers (claim 4). Frieda does not describe the claims' detail of virtual file system, However, APA pages 1-2 describes VFS file are striped across logical storage array. It has been known that any files including meta data files can be distributed to any disks using the file systems such as NFS and VFS.

As for claim 5, Fridella describes wherein a file has attributes and each server for each file maintains a caching element that stores a last known version of the file attributes and ranges of modification time and change time values for assignment to write operation results (Fridella page 4 paragraph 32 describes file attribute are cached in both primary mover and secondary movers; Fridella's page 5 paragraph 38 further describes the secondary movers maintain the latest time using a local value m which obtained from the primary clock and value of a local timer).

As for claim 6, the claim recites wherein each disk element which is not the meta-data server for a virtual file system is an input output secondary. Claim 6 rejected based on the same rationale as in the rejection of claim 4. Fridella further pages 3,4 paragraph 27 describes that the disks are assigned to data movers operating as a front end to the cached disk array.

As for claim 7, Fridella describes wherein ranges of file modification times or file change times are reserved from the meta-data server by the input output secondary (Fridella's pages 4-5 paragraphs 36-38 describes the update time is a function of the clock time obtained from the clock of the primary mover which including a value m obtained from the primary clock).

As for claim 8, Fridella describes wherein the modification and change times in the ranges obtained from the meta-data server are issued to operations already queued at the input output secondary (Fridella's page 5 paragraph 38 describes the secondary movers maintain clock

time and timer for each file that it has opened for asynchronous write access; Fridella's page 5 paragraph 39 further describes the primary file server sends the updated file-modification time to all of the other secondaries).

As for claim 9, the claim recites wherein modification and change times in the ranges obtained from the meta-data server are issued to operations received during a window of time after the ranges are reserved from the meta-data server by the input output secondary (Fridella's page 5 paragraph 38 describes the secondary receives the present value of the primary clock and a value m from the primary clock; Fridella's page 5 paragraph 42 clearly describes the asynchronous write by one client will not be visible to other clients, until "a window of time" when the client issues an NFS commit).

As for claim 10, Fridella describes wherein operations affecting all stripes of a file begin executions first at the meta-data server for a file, and then execute at all input output secondaries, such that operations at the input output secondaries wait only for already executing operations that have already finished their communication with the meta-data server. The claim rejected based on the same rationale as in the rejection of claim 9. Fridella's page 5 paragraphs 40-42 further describe for the attribute changes that affecting other secondaries, the primary server notifies each secondary; therefore the consistency in a file system is maintained.

As for claim 11 the claim recites an apparatus as described in claim 10 wherein operations follow one of at least two locking models, the first of which is to synchronize first with the meta-data server, then begin core execution by synchronizing with other operations executing at the input output secondary, and the second of which is to first synchronize at the meta-data server, and then to synchronize with operations at one or more input output

secondaries that have begun core execution at the input output secondaries. Fridella's page 5 paragraph 38 describes the secondary must obtain the file attributes from the primary fro a first write to a file.

Claims 12-13 rejected under 35 U.S.C. 103(a) as being unpatentable over Fridella et al (US Pub 2005/0044090), APA (US application 10/727169) as applied to claim 11 and further in view of Edsall et al (US 2004/0139167).

As for claims 12-13, the claims recite wherein the cluster network is connected in a star topology (claim 12); wherein the cluster network is a switched Ethernet (claim 13). Although Fridella does not describe the claim's detail of the network. However, Edsell's Fig 1: #22, Fig 5: #204 describes a network utilizing gigabit Ethernet switches to connect plural of servers nodes and disk controller nodes. It would have been obvious to one of ordinary skill in the art at the time of invention to include the switches as suggested by Edsell in Frieda's system to manage and providing concurrent connections among nodes, thereby the data paths among nodes can be dynamically managed and rerouted by the switches in an efficiently manner (Edsell's page 6 paragraph 53).

Claims 14-19,29-35 rejected under 35 U.S.C. 103(a) as being unpatentable over Fridella et al (US Pub 2005/0044090) in view of APA (US application 10/727169) .

As for claim 14, the claim recites creating a file across a plurality of NFS servers; writing data into the file as strips of the data in the servers, the strips together forming a stripe; reading strips of the data from the servers; and deleting the strips from the servers. The claim rejected

based on the same rationale as in the rejection of claim 1. Furthermore, APA's page 2 lines 1-15 describes a PVFS file system in which data is stripped among multiple servers through an additional file system layer built on top of a normal file system. The PVFS file system further provides file access commands to read, write and update data in strip files.

Claim 15 rejected based on the same rationale as in the rejection of claim 6.

Claim 16 rejected based on the same rationale as in the rejection of claim 5.

Claim 17 rejected based on the same rationale as in the rejection of claim 7.

As for claims 18-19, including the step of making a status request by the caching element to the meta-data server to obtain a file's current attributes (claim 18) wherein the making a status request step includes the step of obtaining modification time and change time ranges from the meta-data server (claim 19).

Claims 20, 21,23,25 rejected under 35 U.S.C. 103(a) as being unpatentable over Fridella et al (US Pub 2005/0044090), APA (US application 10/727169) as applied to claim 19 and further in view of Cheng et al (US 5701516).

As for claims 20,21,23,25 the claims recite including the step of queuing file read and file write requests at the input output secondary until the file read and file write requests are admitted by the cache element and complete execution (claim 20); including the step of tracking by the cache element of the file read and file write requests executing for the file and the ranges that are being read or written (claim 21); including the step of checking a byte range affected by a file read request to ensure it does not overlap a byte range of any file write requests previously admitted and currently executing (claim 23); including the step of checking with the cache element the byte range affected by the file write request for overlap with any admitted and still

executing file read or file write requests (claim 25). Fridella's page 5 paragraph 41 taught that in effect, the secondary is requesting a lock on a range of file blocks. Thus Fridella clearly suggest that subsequent requests must be granted strictly in serial order to guarantee the integrity of data being updated, thus the consistency of NFS file system can be achieved. Fridella does not describe the claims' aspect of queuing the requests. However, Cheng describes using a queue for receiving requests is utilized to keep subsequence requests pending while a current request is being executed. When the current request is completed, the subsequent requests can be executed in the order they were received (Cheng's column 17 lines 55-60). It would have been obvious to one of ordinary skill in the art at the time of invention to include the queuing methods and structures as suggested by Cheng in Frieda's system to store and manage multiple requests in a first in first out order, thereby the data accessing to the same disk is executed in a proper order and preserving the disk data integrity (Cheng's column 17 lines 55-60).

As for claims 22,24 the claims recite including the step of requesting the cache element move out of invalid node to read mode when a read operation must be executed (claim 22); including the step of requesting, in response to a file write request that the cache element move into a write mode (claim 24). Fridella's Fig 3 #151 to #159 shows a secondary receives a file access request that moves it into a corresponding state (read or write accessing mode) and executing sequence of steps to fulfill the request.

As for claim 26, Fridella describes including the step, when executing a write request, of allocating a modification time and change time pair from the range of modification times and change times stored in the cache element (Fridella's paragraph 38).

As for claim 27, the claim recites including the step of checking the head of a queue of pending file read and file write requests to see if a head request can be admitted by the caching element after either a file read or file write request is completed. The claim rejected based on the same rationale as in the rejection of claims 20,21,23,25.

As for claim 28 Fridella describes including the steps of detecting by the cache element that a file length must be updated in response to a file write request, moving the cache element into exclusive mode; and making a file write status call to the meta-data server to update length attributes of the file (Fridella's page 4 paragraphs 28,30 describes steps get a file system exclusively and to update file attribute such as file length).

Claim 29 rejected based on the same rationale as in the rejection of claim 5.

Claim 30 rejected based on the same rationale as in the rejection of claim 7.

Claim 31 rejected based on the same rationale as in the rejection of claim 18.

Claim 32 rejected based on the same rationale as in the rejection of claim 19.

Claim 33 rejected based on the same rationale as in the rejection of claim 22.

Claim 34 rejected based on the same rationale as in the rejection of claim 24.

Claim 35 rejected based on the same rationale as in the rejection of claim 28.

Claims 38-39 rejected under 35 U.S.C. 103(a) as being unpatentable over Fridella et al (US Pub 2005/0044090) in view of APA (US application 10/727169) and further in view of Schmuck et al (US 6032216), and Ross (Using the Parallel Virtual File System)).

As for claim 38 the claim recites a method for reading data in a file comprising the steps of: receiving an NFS read request for data in the file at a network element; determining by the network element which VFS stores at least one strip containing the data; sending a file read request from the network element to at least one disk element of a plurality of servers storing a strip of the data; obtaining current attributes associated with the file by each disk element; reading the strips of the file from each disk element having the strips; and generating a response in regard to the file read request. The claim rejected based on the same rationale as in the rejection of claim 37. Fridella does not describes the claim's detail of determining which VFS stores a strip containing the data. However, Ross's page 13 describes a parallel VFS file sytem in which the program using normal Unix I/O, the PVFS system can translate the data access requests into a stripe in a particular disk as shown in page 14, Fig 3. It would have been obvious to one of ordinary skill in the art at the time of invention to include the file accessing methods for PVFS file system as suggested by Ross in Frieda's system to provide translation of normal Unix I/O file accessing commands into proper accessing command of data strip in a PVFS file system. Thereby, user can accesses data in PVFS file system utilizing familiar Unix I/O commands without any changes.

As for claim 39 the claim recites a method for writing data in a file comprising the steps of: receiving an NFS write request for a file at a network element; determining by the network element which VFS is associated with the file; sending a file write request from the network element to at least one disk element of a plurality of servers having a stripe of the VFS; acquiring current attributes associated with the file; and writing a predetermined number of bytes of the

data into each VFS strip in succession until all of the data is written into the file. The claim rejected based on the same rationale as in the rejection of claim 38.

***Response to Arguments***

Applicant's arguments in response to the last office action has been fully considered but they are not persuasive. Examiner respectfully traverses Applicant's arguments for the following reasons:

As to the remarks on pages2-9 concerning the claim 1,

The claim 1 recites "An apparatus for data storage comprising: a cluster of NFS servers, each server having network ports for incoming file system requests and cluster traffic between servers; and a plurality of storage arrays in communication with the servers, the servers utilizing a striped file system for storing data".

A) The claim 1 describes broadly a multiple of NFS servers to handle file system requests and the servers utilizing a striped file system for storing data". These limitations are clearly described by APA in specification pages 2 and 3 that describes a parallel virtual file system PVFS in which data are stripped across multiple disks. The stripping of data in PVFS system (i.e stripping data and accessing data by different clients/servers paths) is well known in the art. In fact, APA describes the stripping data blocks are maintained by a structure using information of file systems in PVFS (APA's column 2 lines 1-3). In fact, APA's PVFS handles requests of NFS file system type by using a locking mechanism (specification's page 2 last line to page 3 line 1) , thus guaranteeing the integrity of the file and the correctness of the file's modification time. Furthermore, the specification pages 2 and 3 clearly describes the invention differs from the current art mainly by providing a more efficiently mechanism to generate a file's

modification time when the file is being modified. Examiner further notes that none of the features regarding the file's modification time are recited in claim 1. Therefore, APA in specifications pages 2 and 3 clearly suggests all limitations of claim 1.

B) In a similar manner, Fridella describes a parallel virtual file system PVFS having multiple servers **to execute multiple data requests of a virtual file system**, wherein the access to data of a file in a virtual file system is done using any server nodes and their file systems **in a parallel manner.** Fridella further describes the file system being NFS file system type (Fig 1, paragraph 7). Thus Fridella clearly suggests that the system is capable of executing multiple data requests for stripped data blocks, since these data block readily to be accessed in a parallel manner. One skilled in the art would have been motivated to use stripped data in Fridella's system, since by stripping data blocks, these data are accessed parallel using multiple servers paths as taught by Fridella and thereby the overall throughput of the system would be increased. Fridella further describes a mechanism that efficiently guarantees the integrity of the file and the correctness of the file's modification time (see Fridella's abstract).

Therefor the teachings of Fridella and APA is deemed to meet the claim's limitation.

C) Per page 7 in the remark, applicant argues that "different servers having different clocks..with the result that later writes do not advance a systemwide file modification time". Fridella describes a mechanism that for the first write request (for example first data request of data in each strip, in each server), the secondary server obtains the meta data information of this write request from the primary server (corresponding to applicant's meta data server). The meta data information includes the clock time of the meta data server. Because every first data

modification request uses the same clock source (from the meta data server). The secondary server is only allowed to increment the clock for subsequent write requests at its node. Examiner notes that specification's page 25 describes a similar mechanism. For example, the "meta data " server I0 informs the "secondary" server I2 that the arrival time of the first write request is at 200000, and for the fifty subsequent writes at the "secondary server" I2, it simply adds 50 units of request time. Again, there is only one source of "meta data" server that gives the time/clock information to the secondary server".

D) Per remarks on page 8-10 for claim 13, the claim merely recite a network connects by a switch and using the switch device to provide multiple paths from servers to storage devices. Examiner maintains that Edsell teaches the limitation in the claim.

E) Per the remarks on pages 10-11 for claims 20-23,25. Again, the stripping of data has been discussed in above paragraphs. Regardless of the stripping of data or not, it has been known in the art that a queue to hold the requests being received are normally provide to buffering the pending requests while waiting for the current request complete its execution. Subsequently, the pending requests will be executed in a serial manner as they were obtained fr4om the queue. Cheng describes such queue in his disclosure.

F) Per the remarks on page 12 for claims 38-39. Examiner maintains that Ross clearly teaches a mechanism that track the host i/o access request in a PVFS system, translating to proper data i/o requests for stripping data across each server in the PVFS system. (see Ross' Fig 3).

***Allowable Subject Matter***

Claims 36,37 are allowed.

*Conclusion*

**THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

When responding to the office action, Applicant is advised to provide the examiner with the line numbers and page numbers in the application and/or references cited to assist examiner to locate the appropriate paragraphs.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Duc T. Doan whose telephone number is 571-272-4171. The examiner can normally be reached on M-F 8:00 AM 05:00 PM.

- If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Mano Padmanabhan can be reached on 571-272-4210. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Mano Padmanabhan
5/1/06

MANO PADMANABHAN
SUPERVISORY PATENT EXAMINER